



PCT

RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/10/505,263

TIME: 15:58:27

Input Set : D:\1242-50-2-PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505263.raw

3 <110> APPLICANT: Vanderbilt University
 4 Case Western Reserve University
 5 The Brigham and Women's Hospital, Inc.
 6 Mount, David B
 7 Romero, Michael
 9 <120> TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and
 SLC26A2
 10 ANION EXCHANGERS
 12 <130> FILE REFERENCE: 1242/50/2 PCT/US
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/505,263
 C--> 14 <141> CURRENT FILING DATE: 2004-08-20
 14 <150> PRIOR APPLICATION NUMBER: US 60/360,275
 15 <151> PRIOR FILING DATE: 2002-02-28
 17 <150> PRIOR APPLICATION NUMBER: PCT/US03/06469
 18 <151> PRIOR FILING DATE: 2003-02-28
 20 <160> NUMBER OF SEQ ID NOS: 96
 22 <170> SOFTWARE: PatentIn version 3.2
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 2654
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Homo sapiens
 29 <400> SEQUENCE: 1
 30 acgcgtccgc ctctagctcg agcagcagga gcagcccgca ccggacaact tgcgagccat 60
 32 ggggctggcg gatgcgtcgg gaccgaggga cacacaggca ctgctgtctg caacacaagc 120
 34 aatggacctg cggaggcgag actaccacat ggaacggccg ctgctgaacc aggagcattt 180
 36 ggaggagctg gggcgctggg gctcagcacc taggaccac cagtggcgga cctgggttga 240
 38 gtgctcccggt gctcgggctt atgcccttct gctccaacac ctcccggttt tgggtctggtt 300
 40 accccgggtat cctgtgcgtg actggtcctt ggggtgacctg ttatccggcc tgagtgtggc 360
 42 catcatgcag cttccgcagg gcttggccta cgccctcctg gctggattgc ccccgtggtt 420
 44 tggcctctat agctccttct accctgtctt catctacttc ctgtttggca cttcccggca 480
 46 catctccgtg gggacctttg ctgtcatgtc tgtgatggtg ggcagtgtga cagaatccct 540
 48 ggccccgcag gccttgaacg actccatgat caatgagaca gccagagatg ctgcccgggt 600
 50 acaggtggcc tccacactca gtgtcctggt tggcctcttc caggtggggc tgggcctgat 660
 52 ccacttcggc ttcgtggtea cctacctgtc agaacctctt gtccgaggct ataccacagc 720
 54 tgcagctgtg caggtcttcg tctcacagct caagtatgtg tttggcctcc atctgagcag 780
 56 ccactctggg ccactgtccc tcatctatac agtgcgtggg gtctgctgga agctgccccca 840
 58 gagcaagggtt ggcaccgtgg tcaactgcagc tgtggctggg gtggtgctcg tgggtggtgaa 900
 60 gctgttgaat gacaagctgc agcagcagct gcccatgccg ataccggggg agctgctcac 960
 62 gctcatcggg gccacaggca tctcctatgg catgggtcta aagcacagat ttgaggtaga 1020
 64 tgtcgtgggc aacatccctg cagggtcgtt gcccccagtg gcccccaaca cccagctggt 1080
 66 ctcaaagctc gtgggcagcg ccttcaccat cgctgtggtt ggggttgcca ttgccatctc 1140
 68 actggggaag atcttcgcgc tgaggcacgg ctaccgggtg gacagcaacc aggagctggt 1200
 70 ggccctgggc cttcagtaacc ttatcggagg catcttcacg tgcttccccg tgagttgctc 1260
 72 tatgtctcgg agcctggtac aggagagcac cgggggcaac tcgcaggttg ctggagccat 1320

(pg. 6)

RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/10/505,263

TIME: 15:58:27

Input Set : D:\1242-50-2-PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505263.raw

```

74 ctcttccctt ttcattcctcc tcatcattgt caaacttggg gaactcttcc atgacctgcc 1380
76 caaggcgggc ctggcagcca tcatcattgt gaacctgaag ggcattgctga ggcagctcag 1440
78 cgacatgcgc tccctctgga aggccaatcg ggcggatctg cttatctggc tggtagacctt 1500
80 caccggccacc atcttgctga acctggacct tggcttggtg gttgcgggtca tcttctccct 1560
82 gctgctcgtg gtggtccgga cacagatgcc ccactactct gtcttggggc aggtgccaga 1620
84 caccgatatt tacagagatg tggcagagta ctcagaggcc aagggaagtcc ggggggtgaa 1680
86 ggtcttccgc tcctcggcca ccgtgtactt tgccaatgct gagttctaca gtgatgcgct 1740
88 gaagcagagg tgtggtgtgg atgtcgactt cctcatctcc cagaagaaga aactgctcaa 1800
90 gaagcaggag cagctgaagc tgaagcaact gcagaaagag gagaagcttc ggaaacaggc 1860
92 tgcttcccc aagggcgcct cagtttccat taatgtcaac accagccttg aagacatgag 1920
94 gagcaacaac gttgaggact gcaagatgat ggtgagctca ggagataaga tggagatgc 1980
96 aacagccaat ggtcaagaag actccaaggc cccagatggg tccacactga aggccttggg 2040
98 cctgcctcag ccagacttcc acagcctcat cctggacctg ggtgcctct cctttgtgga 2100
100 cactgtgtgc ctcaagagcc tgaagaatat ttccatgac ttccgggaga ttgaggtgga 2160
102 ggtgtacatg gcggcctgcc acagcctgt ggtcagccag cttgaggctg ggcacttctt 2220
104 cgatgcattc atcaccaaga agcatctctt tgctctgtc catgatgctg tcacctttgc 2280
106 cctccaacac ccgaggcctg tccccgacag cctgtttcg gtcaccagac tctgaacatg 2340
108 ctacatcctg cccaagactg cacctctgga ggtgcagggc acccttgaga agcccctcac 2400
110 ccctaggccg cctccagggt ctaccagga gtccctcca tgtacacaca cacaactcag 2460
112 ggaaggaggt cctgggactc caagttcagc gctccagggtc tgggacaggc cctgcatgca 2520
114 gtcaggctgg cagtggcgcg gtacaggag ggaactggtg catattttag cctcaggaat 2580
116 aaagatttgt ctgcaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2640
118 aaaaaaaaaa aaaa 2654
121 <210> SEQ ID NO: 2
122 <211> LENGTH: 758
123 <212> TYPE: PRT
124 <213> ORGANISM: Homo sapiens
126 <400> SEQUENCE: 2
128 Met Gly Leu Ala Asp Ala Ser Gly Pro Arg Asp Thr Gln Ala Leu Leu
129 1 5 10 15
132 Ser Ala Thr Gln Ala Met Asp Leu Arg Arg Arg Asp Tyr His Met Glu
133 20 25 30
136 Arg Pro Leu Leu Asn Gln Glu His Leu Glu Glu Leu Gly Arg Trp Gly
137 35 40 45
140 Ser Ala Pro Arg Thr His Gln Trp Arg Thr Trp Leu Gln Cys Ser Arg
141 50 55 60
144 Ala Arg Ala Tyr Ala Leu Leu Leu Gln His Leu Pro Val Leu Val Trp
145 65 70 75 80
148 Leu Pro Arg Tyr Pro Val Arg Asp Trp Leu Leu Gly Asp Leu Leu Ser
149 85 90 95
152 Gly Leu Ser Val Ala Ile Met Gln Leu Pro Gln Gly Leu Ala Tyr Ala
153 100 105 110
156 Leu Leu Ala Gly Leu Pro Pro Val Phe Gly Leu Tyr Ser Ser Phe Tyr
157 115 120 125
160 Pro Val Phe Ile Tyr Phe Leu Phe Gly Thr Ser Arg His Ile Ser Val
161 130 135 140
164 Gly Thr Phe Ala Val Met Ser Val Met Val Gly Ser Val Thr Glu Ser
165 145 150 155 160
168 Leu Ala Pro Gln Ala Leu Asn Asp Ser Met Ile Asn Glu Thr Ala Arg

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/505,263

DATE: 08/26/2004

TIME: 15:58:27

Input Set : D:\1242-50-2-PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505263.raw

169		165		170		175
172	Asp Ala Ala Arg Val Gln Val Ala Ser Thr Leu Ser Val Leu Val Gly					
173		180		185		190
176	Leu Phe Gln Val Gly Leu Gly Leu Ile His Phe Gly Phe Val Val Thr					
177		195		200		205
180	Tyr Leu Ser Glu Pro Leu Val Arg Gly Tyr Thr Thr Ala Ala Ala Val					
181		210		215		220
184	Gln Val Phe Val Ser Gln Leu Lys Tyr Val Phe Gly Leu His Leu Ser					
185	225		230		235	240
188	Ser His Ser Gly Pro Leu Ser Leu Ile Tyr Thr Val Leu Glu Val Cys					
189		245		250		255
192	Trp Lys Leu Pro Gln Ser Lys Val Gly Thr Val Val Thr Ala Ala Val					
193		260		265		270
196	Ala Gly Val Val Leu Val Val Val Lys Leu Leu Asn Asp Lys Leu Gln					
197		275		280		285
200	Gln Gln Leu Pro Met Pro Ile Pro Gly Glu Leu Leu Thr Leu Ile Gly					
201		290		295		300
204	Ala Thr Gly Ile Ser Tyr Gly Met Gly Leu Lys His Arg Phe Glu Val					
205	305		310		315	320
208	Asp Val Val Gly Asn Ile Pro Ala Gly Leu Val Pro Pro Val Ala Pro					
209		325		330		335
212	Asn Thr Gln Leu Phe Ser Lys Leu Val Gly Ser Ala Phe Thr Ile Ala					
213		340		345		350
216	Val Val Gly Phe Ala Ile Ala Ile Ser Leu Gly Lys Ile Phe Ala Leu					
217		355		360		365
220	Arg His Gly Tyr Arg Val Asp Ser Asn Gln Glu Leu Val Ala Leu Gly					
221		370		375		380
224	Leu Ser Asn Leu Ile Gly Gly Ile Phe Gln Cys Phe Pro Val Ser Cys					
225	385		390		395	400
228	Ser Met Ser Arg Ser Leu Val Gln Glu Ser Thr Gly Gly Asn Ser Gln					
229		405		410		415
232	Val Ala Gly Ala Ile Ser Ser Leu Phe Ile Leu Leu Ile Ile Val Lys					
233		420		425		430
236	Leu Gly Glu Leu Phe His Asp Leu Pro Lys Ala Val Leu Ala Ala Ile					
237		435		440		445
240	Ile Ile Val Asn Leu Lys Gly Met Leu Arg Gln Leu Ser Asp Met Arg					
241		450		455		460
244	Ser Leu Trp Lys Ala Asn Arg Ala Asp Leu Leu Ile Trp Leu Val Thr					
245	465		470		475	480
248	Phe Thr Ala Thr Ile Leu Leu Asn Leu Asp Leu Gly Leu Val Val Ala					
249		485		490		495
252	Val Ile Phe Ser Leu Leu Leu Val Val Val Arg Thr Gln Met Pro His					
253		500		505		510
256	Tyr Ser Val Leu Gly Gln Val Pro Asp Thr Asp Ile Tyr Arg Asp Val					
257		515		520		525
260	Ala Glu Tyr Ser Glu Ala Lys Glu Val Arg Gly Val Lys Val Phe Arg					
261		530		535		540
264	Ser Ser Ala Thr Val Tyr Phe Ala Asn Ala Glu Phe Tyr Ser Asp Ala					
265	545		550		555	560

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/505,263

DATE: 08/26/2004

TIME: 15:58:27

Input Set : D:\1242-50-2-PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505263.raw

```

268 Leu Lys Gln Arg Cys Gly Val Asp Val Asp Phe Leu Ile Ser Gln Lys
269           565           570           575
272 Lys Lys Leu Leu Lys Lys Gln Glu Gln Leu Lys Leu Lys Gln Leu Gln
273           580           585           590
276 Lys Glu Glu Lys Leu Arg Lys Gln Ala Ala Ser Pro Lys Gly Ala Ser
277           595           600           605
280 Val Ser Ile Asn Val Asn Thr Ser Leu Glu Asp Met Arg Ser Asn Asn
281           610           615           620
284 Val Glu Asp Cys Lys Met Met Val Ser Ser Gly Asp Lys Met Glu Asp
285 625           630           635           640
288 Ala Thr Ala Asn Gly Gln Glu Asp Ser Lys Ala Pro Asp Gly Ser Thr
289           645           650           655
292 Leu Lys Ala Leu Gly Leu Pro Gln Pro Asp Phe His Ser Leu Ile Leu
293           660           665           670
296 Asp Leu Gly Ala Leu Ser Phe Val Asp Thr Val Cys Leu Lys Ser Leu
297           675           680           685
300 Lys Asn Ile Phe His Asp Phe Arg Glu Ile Glu Val Glu Val Tyr Met
301           690           695           700
304 Ala Ala Cys His Ser Pro Val Val Ser Gln Leu Glu Ala Gly His Phe
305 705           710           715           720
308 Phe Asp Ala Ser Ile Thr Lys Lys His Leu Phe Ala Ser Val His Asp
309           725           730           735
312 Ala Val Thr Phe Ala Leu Gln His Pro Arg Pro Val Pro Asp Ser Pro
313           740           745           750
316 Val Ser Val Thr Arg Leu
317           755
320 <210> SEQ ID NO: 3
321 <211> LENGTH: 2748
322 <212> TYPE: DNA
323 <213> ORGANISM: Homo sapiens
326 <220> FEATURE:
327 <221> NAME/KEY: CDS
328 <222> LOCATION: (252)..(2468)
330 <400> SEQUENCE: 3
331 aaagaggctt tctcaccagc caaagttcag gtactgcccc tgtctgagtt tggacatgtg      60
333 gccacagat gccagctggg ggctggaata ctgtaccttg atcttagggc ctggaaaaat      120
335 caggtaggta actctctggt cctaggtgga ggtgtagggg aggttgggcc agggatgcct      180
337 tcaactgtgtc tctctggtct tgccacccca gaccgagggg cacacaggca ctgctgtctg      240
339 caacacaagc a atg gac ctg cgg agg cga gac tac cac atg gaa cgg ccg      290
340           Met Asp Leu Arg Arg Arg Asp Tyr His Met Glu Arg Pro
341           1           5           10
343 ctg ctg aac cag gag cat ttg gag gag ctg ggg cgc tgg ggc tca gca      338
344 Leu Leu Asn Gln Glu His Leu Glu Glu Leu Gly Arg Trp Gly Ser Ala
345           15           20           25
347 cct agg acc cac cag tgg cgg acc tgg ttg cag tgc tcc cgt gct cgg      386
348 Pro Arg Thr His Gln Trp Arg Thr Trp Leu Gln Cys Ser Arg Ala Arg
349 30           35           40           45
351 gcc tat gcc ctt ctg ctc caa cac ctc ccg gtt ttg gtc tgg tta ccc      434
352 Ala Tyr Ala Leu Leu Leu Gln His Leu Pro Val Leu Val Trp Leu Pro

```

RAW SEQUENCE LISTING

DATE: 08/26/2004

PATENT APPLICATION: US/10/505,263

TIME: 15:58:27

Input Set : D:\1242-50-2-PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505263.raw

353		50		55		60	
355	cgg tat cct gtg cgt gac tgg ctc ctg ggt gac ctg tta tcc ggc ctg						482
356	Arg Tyr Pro Val Arg Asp Trp Leu Leu Gly Asp Leu Leu Ser Gly Leu						
357		65		70		75	
359	agt gtg gcc atc atg cag ctt ccg cag ggc ttg gcc tac gcc ctc ctg						530
360	Ser Val Ala Ile Met Gln Leu Pro Gln Gly Leu Ala Tyr Ala Leu Leu						
361		80		85		90	
363	gct gga ttg ccc ccc gtg ttt ggc ctc tat agc tcc ttc tac cct gtc						578
364	Ala Gly Leu Pro Pro Val Phe Gly Leu Tyr Ser Ser Phe Tyr Pro Val						
365		95		100		105	
367	ttc atc tac ttc ctg ttt ggc act tcc cgg cac atc tcc gtg ggg acc						626
368	Phe Ile Tyr Phe Leu Phe Gly Thr Ser Arg His Ile Ser Val Gly Thr						
369	110		115		120		125
371	ttt gct gtc atg tct gtg atg gtg ggc agt gtg aca gaa tcc ctg gcc						674
372	Phe Ala Val Met Ser Val Met Val Gly Ser Val Thr Glu Ser Leu Ala						
373		130		135		140	
375	ccg cag gcc ttg aac gac tcc atg atc aat gag aca gcc aga gat gct						722
376	Pro Gln Ala Leu Asn Asp Ser Met Ile Asn Glu Thr Ala Arg Asp Ala						
377		145		150		155	
379	gcc cgg gta cag gtg gcc tcc aca ctc agt gtc ctg gtt ggc ctc ttc						770
380	Ala Arg Val Gln Val Ala Ser Thr Leu Ser Val Leu Val Gly Leu Phe						
381		160		165		170	
383	cag gtg ggg ctg ggc ctg atc cac ttc ggc ttc gtg gtc acc tac ctg						818
384	Gln Val Gly Leu Gly Leu Ile His Phe Gly Phe Val Val Thr Tyr Leu						
385		175		180		185	
387	tca gaa cct ctt gtc cga ggc tat acc aca gct gca gct gtg cag gtc						866
388	Ser Glu Pro Leu Val Arg Gly Tyr Thr Thr Ala Ala Ala Val Gln Val						
389	190		195		200		205
391	ttc gtc tca cag ctc aag tat gtg ttt ggc ctc cat ctg agc agc cac						914
392	Phe Val Ser Gln Leu Lys Tyr Val Phe Gly Leu His Leu Ser Ser His						
393		210		215		220	
395	tct ggg cca ctg tcc ctc atc tat aca gtg ctg gag gtc tgc tgg aag						962
396	Ser Gly Pro Leu Ser Leu Ile Tyr Thr Val Leu Glu Val Cys Trp Lys						
397		225		230		235	
399	ctg ccc cag agc aag gtt ggc acc gtg gtc act gca gct gtg gct ggg						1010
400	Leu Pro Gln Ser Lys Val Gly Thr Val Val Thr Ala Ala Val Ala Gly						
401		240		245		250	
403	gtg gtg ctc gtg gtg gtg aag ctg ttg aat gac aag ctg cag cag cag						1058
404	Val Val Leu Val Val Val Lys Leu Leu Asn Asp Lys Leu Gln Gln Gln						
405		255		260		265	
407	ctg ccc atg ccg ata ccc ggg gag ctg ctc acg ctc atc ggg gcc aca						1106
408	Leu Pro Met Pro Ile Pro Gly Glu Leu Leu Thr Leu Ile Gly Ala Thr						
409	270		275		280		285
411	ggc atc tcc tat ggc atg ggt cta aag cac aga ttt gag gta gat gtc						1154
412	Gly Ile Ser Tyr Gly Met Gly Leu Lys His Arg Phe Glu Val Asp Val						
413		290		295		300	
415	gtg ggc aac atc cct gca ggg ctg gtg ccc cca gtg gcc ccc aac acc						1202
416	Val Gly Asn Ile Pro Ala Gly Leu Val Pro Pro Val Ala Pro Asn Thr						
417		305		310		315	

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/505,263

DATE: 08/26/2004
TIME: 15:58:28

Input Set : D:\1242-50-2-PCT-US.ST25.txt
Output Set: N:\CRF4\08262004\J505263.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; N Pos. 2850

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/505,263

DATE: 08/26/2004

TIME: 15:58:28

Input Set : D:\1242-50-2-PCT-US.ST25.txt

Output Set: N:\CRF4\08262004\J505263.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:1423 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:2820